II. RAYS: THE EIKONAL TREATMENT OF GEOMETRIC OPTICS [6]

Since ancient times, the notion of **ray or beam propagation** has been one of the most enduring and fundamental concepts in optical physics. As a zeroth order approximation we might consider a plane wave to be a model of a **beam** and its propagation vector to be a model of a **ray**. This is a reasonable start, but it is a much too restricted view and we can do much better. What we need is a solution to Maxwell's equations which is like a plane wave, but limited in spatial extent. One approach, the simplest, is called variously ray, Gaussian or geometric optics.

A MAXWELLIAN DERIVATION OF THE EIKONAL EQUATION:

To fully understand geometric optics in the context of Maxwell's equations, we start by writing the electric and magnetic fields as ***pseudo-simple*** waves -- *viz.*

$$\vec{\mathbf{E}}\left(\vec{\mathbf{r}},\omega\right) = \vec{\mathbf{e}}\left(\vec{\mathbf{r}},\omega\right)\,\exp\left[-i\,\mathrm{k}_0\,\mathrm{S}\left(\vec{\mathbf{r}},\omega\right)\right] \qquad \text{[II-1a]}$$

$$\vec{\mathbf{H}}\left(\vec{\mathbf{r}},\omega\right) = \vec{\mathbf{h}}\left(\vec{\mathbf{r}},\omega\right)\,\exp\left[-i\,\mathrm{k}_0\,\mathrm{S}\left(\vec{\mathbf{r}},\omega\right)\right] \qquad \text{[II-1b]}$$

where $\mathrm{k}_0 = \omega\sqrt{\mu_0\varepsilon_0} = \omega / c$

It is assumed that $\vec{\mathbf{e}}\left(\vec{\mathbf{r}},\omega\right)$ and $\vec{\mathbf{h}}\left(\vec{\mathbf{r}},\omega\right)$ are **weak functions of position**. The scalar phase function $\mathrm{S}\left(\vec{\mathbf{r}},\omega\right)$ is the spatially varying phase of the *pseudo-simple* wave. For the cases of pseudo-plane waves and pseudo-spherical waves the phase function is given, respectively, by

$$\mathrm{k}_0\mathrm{S}\left(\vec{\mathbf{r}},\omega\right) = x\,\mathrm{k}_x + y\,\mathrm{k}_y + z\,\mathrm{k}_z \qquad \text{[II-2a]}$$

[6] See, for example, Max Born and Emil Wolf, *Principle of Optics*, Pergamon Press (1986), Chapter 3.

and $$k_0 S(\vec{\mathbf{r}},\omega) = k_0 \sqrt{x^2 + y^2 + z^2} \qquad \text{[II-2b]}$$

We now substitute these pseudo-simple wave expressions (*i.e.* Equations [II-1]) into Maxwell's equations to obtain

$$\exp\left[-i k_0 S(\vec{\mathbf{r}},\omega)\right]\left\{ \; \times\vec{\mathbf{e}}(\vec{\mathbf{r}},\omega) - i k_0 \; S(\vec{\mathbf{r}},\omega)\times\vec{\mathbf{e}}(\vec{\mathbf{r}},\omega)\right\} = -i\mu_0 c k_0 \vec{\mathbf{h}}(\vec{\mathbf{r}},\omega)\exp\left[-i k_0 S(\vec{\mathbf{r}},\omega)\right] \qquad [$$

$$\exp\left[-i k_0 S(\vec{\mathbf{r}},\omega)\right]\left\{ \; \times\vec{\mathbf{h}}(\vec{\mathbf{r}},\omega) - i k_0 \; S(\vec{\mathbf{r}},\omega)\times\vec{\mathbf{h}}(\vec{\mathbf{r}},\omega)\right\} = i\varepsilon(\vec{\mathbf{r}},\omega) c k_0 \vec{\mathbf{e}}(\vec{\mathbf{r}},\omega)\exp\left[-i k_0 S(\vec{\mathbf{r}},\omega)\right] \qquad [$$

$$\exp\left[-i k_0 S(\vec{\mathbf{r}},\omega)\right]\left\{ \; \left[\varepsilon(\vec{\mathbf{r}},\omega)\vec{\mathbf{e}}(\vec{\mathbf{r}},\omega)\right] - i k_0 \varepsilon(\vec{\mathbf{r}},\omega) \; S(\vec{\mathbf{r}},\omega) \; \vec{\mathbf{e}}(\vec{\mathbf{r}},\omega)\right\} = 0 \qquad [$$

$$\exp\left[-i k_0 S(\vec{\mathbf{r}},\omega)\right]\left\{ \; \vec{\mathbf{h}}(\vec{\mathbf{r}},\omega) - i k_0 \; S(\vec{\mathbf{r}},\omega) \; \vec{\mathbf{h}}(\vec{\mathbf{r}},\omega)\right\} = 0 \qquad [$$

Rearranging, we obtain

$$S(\vec{\mathbf{r}},\omega)\times\vec{\mathbf{e}}(\vec{\mathbf{r}},\omega) - \mu_0 c\, \vec{\mathbf{h}}(\vec{\mathbf{r}},\omega) = [i k_0]^{-1} \; \times\vec{\mathbf{e}}(\vec{\mathbf{r}},\omega) \qquad \text{[II-4a]}$$

$$S(\vec{\mathbf{r}},\omega)\times\vec{\mathbf{h}}(\vec{\mathbf{r}},\omega) + \varepsilon(\vec{\mathbf{r}},\omega) c\,\vec{\mathbf{e}}(\vec{\mathbf{r}},\omega) = [i k_0]^{-1} \; \times\vec{\mathbf{h}}(\vec{\mathbf{r}},\omega) \qquad \text{[II-4b]}$$

$$S(\vec{\mathbf{r}},\omega) \; \left[\varepsilon(\vec{\mathbf{r}},\omega)\vec{\mathbf{e}}(\vec{\mathbf{r}},\omega)\right] = [i k_0]^{-1} \; \left[\varepsilon(\vec{\mathbf{r}},\omega)\vec{\mathbf{e}}(\vec{\mathbf{r}},\omega)\right] \qquad \text{[II-4c]}$$

$$S(\vec{\mathbf{r}},\omega) \; \vec{\mathbf{h}}(\vec{\mathbf{r}},\omega) = [i k_0]^{-1} \; \vec{\mathbf{h}}(\vec{\mathbf{r}},\omega) \qquad \text{[II-4d]}$$

In the **ray**, **Gaussian** or **geometric approximation** we assume that we may neglect the RHS's of these equations. To get something useful we multiply through the first equation (*i.e.* Equation [II-4a]) as follows:

$$[\mu_0 c]^{-1} \; S(\vec{\mathbf{r}},\omega)\times\{\text{ Equation [II-4a] }\} \qquad \text{[II-5a]}$$

$$[\mu_0 c]^{-1}\,\vec{\;}\,\mathrm{S}(\vec{\mathbf{r}},\omega)\times\left\{\vec{\;}\,\mathrm{S}(\vec{\mathbf{r}},\omega)\times\vec{\mathbf{e}}(\vec{\mathbf{r}},\omega)-\mu_0\,\mathrm{c}\,\vec{\mathbf{h}}(\vec{\mathbf{r}},\omega)\right\}\approx 0 \qquad \text{[II-5b]}$$

Applying the "abc = bac - cab" rule[7] we obtain

$$\begin{aligned}[\mu_0 c]^{-1}\Big\{\,\vec{\;}\,\mathrm{S}(\vec{\mathbf{r}},\omega)\left[\vec{\;}\,\mathrm{S}(\vec{\mathbf{r}},\omega)\;\;\vec{\mathbf{e}}(\vec{\mathbf{r}},\omega)\right]&\\ -\vec{\mathbf{e}}(\vec{\mathbf{r}},\omega)\left|\vec{\;}\,\mathrm{S}(\vec{\mathbf{r}},\omega)\right|^2-\mu_0\,\mathrm{c}\,\vec{\;}\,\mathrm{S}(\vec{\mathbf{r}},\omega)\times\vec{\mathbf{h}}(\vec{\mathbf{r}},\omega)\Big\}&\approx 0\end{aligned} \qquad \text{[II-5c]}$$

which becomes upon substitution from the second Equation [II-4b]

$$\begin{aligned}[\mu_0 c]^{-1}\Big\{\,\vec{\;}\,\mathrm{S}(\vec{\mathbf{r}},\omega)\left[\vec{\;}\,\mathrm{S}(\vec{\mathbf{r}},\omega)\;\;\vec{\mathbf{e}}(\vec{\mathbf{r}},\omega)\right]&\\ -\vec{\mathbf{e}}(\vec{\mathbf{r}},\omega)\left|\vec{\;}\,\mathrm{S}(\vec{\mathbf{r}},\omega)\right|^2\Big\}+\varepsilon(\vec{\mathbf{r}},\omega)\,\mathrm{c}\,\vec{\mathbf{e}}(\vec{\mathbf{r}},\omega)&\approx 0\end{aligned} \qquad \text{[II-5d]}$$

From Equation [II-4c], we see that the first term vanishes in the geometric approximation -- *i.e.*, if we neglect the term $[i\,\mathrm{k}_0]^{-1}\,\vec{\;}\left[\,\varepsilon(\vec{\mathbf{r}},\omega)\;\vec{\mathbf{e}}(\vec{\mathbf{r}},\omega)\right]$. Therefore, for non-vanishing $\vec{\mathbf{e}}(\vec{\mathbf{r}},\omega)$ we obtain the following ***reduction*** of Maxwell's equations:

$$\left|\vec{\;}\,\mathrm{S}(\vec{\mathbf{r}},\omega)\right|^2\approx\varepsilon(\vec{\mathbf{r}},\omega)\,\mu_0\,c^2=n^2(\vec{\mathbf{r}},\omega) \qquad \text{[II-6]}$$

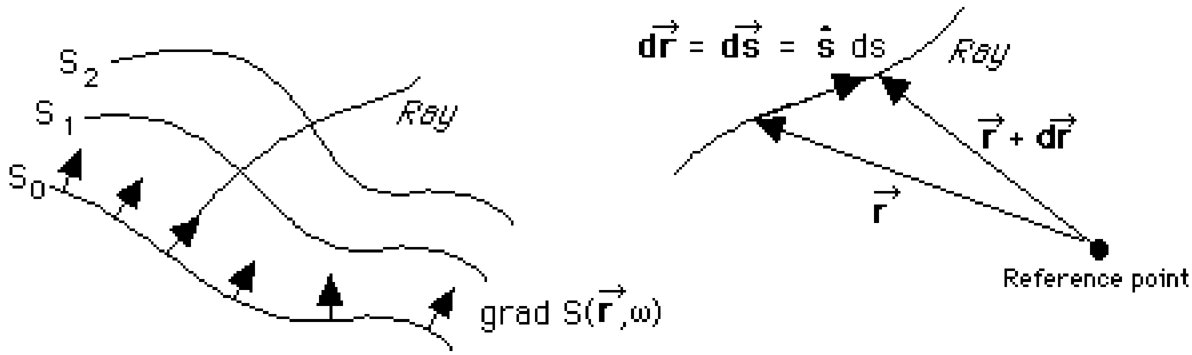
where $n(\vec{\mathbf{r}},\omega)$ is the index of refraction. More explicitly, we may write an equation for a "ray vector" -- *i.e.* the tangent to a space curve orthogonal to the surfaces of constant $\mathrm{S}(\vec{\mathbf{r}},\omega)$

---

7 Again using $\vec{\mathbf{a}}\times(\vec{\mathbf{b}}\times\vec{\mathbf{c}})=\vec{\mathbf{b}}(\vec{\mathbf{a}}\;\;\vec{\mathbf{c}})-\vec{\mathbf{c}}(\vec{\mathbf{a}}\;\;\vec{\mathbf{b}})$.

$$\vec{\nabla} S(\vec{\mathbf{r}},\omega) = n(\vec{\mathbf{r}},\omega)\ \hat{\mathbf{s}} = n(\vec{\mathbf{r}},\omega)\ \frac{d\vec{\mathbf{r}}}{ds} \qquad \text{[II-7]}$$

We illustrate the geometric relationships below:



We may now derive the all important **eikonal equation.** To that end, we first take a derivative along the ray direction -- *viz*.

$$\frac{d}{ds}\left[\vec{\nabla} S(\vec{\mathbf{r}},\omega)\right] = \frac{d}{ds}\left[n(\vec{\mathbf{r}},\omega)\ \frac{d\vec{\mathbf{r}}}{ds}\right] \qquad \text{[II-8a]}$$

However, from the definition of the **grad** operator we know that

$$d\left[\vec{\nabla} S(\vec{\mathbf{r}},\omega)\right] = d\vec{\mathbf{r}} \cdot \vec{\nabla}\left[\vec{\nabla} S(\vec{\mathbf{r}},\omega)\right]$$

so that

$$\frac{d\vec{\mathbf{r}}}{ds} \cdot \vec{\nabla}\left[\vec{\nabla} S(\vec{\mathbf{r}},\omega)\right] = \left\{\frac{1}{n(\vec{\mathbf{r}},\omega)}\ \vec{\nabla} S(\vec{\mathbf{r}},\omega)\right\} \cdot \vec{\nabla}\left[\vec{\nabla} S(\vec{\mathbf{r}},\omega)\right] = \frac{d}{ds}\left[n(\vec{\mathbf{r}},\omega)\ \frac{d\vec{\mathbf{r}}}{ds}\right] \qquad \text{[II-8b]}$$

or

$$\frac{1}{2\,n\left(\vec{\mathbf{r}},\omega\right)}\,\vec{\nabla}\left\{\vec{\nabla}S\left(\vec{\mathbf{r}},\omega\right)\cdot\vec{\nabla}S\left(\vec{\mathbf{r}},\omega\right)\right\}=\frac{1}{2\,n\left(\vec{\mathbf{r}},\omega\right)}\,\vec{\nabla}\left\{n^{2}\left(\vec{\mathbf{r}},\omega\right)\right\}=\vec{\nabla}\,n\left(\vec{\mathbf{r}},\omega\right)=\frac{d}{ds}\left[n\left(\vec{\mathbf{r}},\omega\right)\frac{d\vec{\mathbf{r}}}{ds}\right] \qquad \text{[II-8c]}$$

Thus we have obtain the ***eikonal*** [8] equation for the ray vector -- *viz*.

$$\frac{d}{ds}\left[n\left(\vec{\mathbf{r}},\omega\right)\frac{d\vec{\mathbf{r}}}{ds}\right]=\vec{\nabla}\,n\left(\vec{\mathbf{r}},\omega\right) \qquad \text{[II-9]}$$

FIRST APPLICATION OF THE EIKONAL EQUATION: MIRAGES

Air adjacent to a hot surface rises in temperature and becomes less dense. Thus over a flat hot surface, such as a desert expanse or a sun drenched roadway, air density **locally** increases with height and the average **refractive index** may be approximated by a simple linear variation of the form

$$n\,(x)=n_{g}\,\{1+\kappa\,x\} \qquad \text{[II-10]}$$

where x is the vertical height above the planar surface, n_g is the refractive index at ground level, and κ is a positive constant.

We may use the **eikonal equation** to find an equation for the approximate ray trajectory -- *i.e.* an equation for ray height x as a function of ground distance z -- of a light ray launched from a height x_o and at an angle θ_o with respect to the surface of the earth.

---

8 The *eikonal* (from the Greek: εικϖν means *image*) was introduced in 1895 by H. Bruns.

Therefore,

$$\frac{d}{ds}\left[n(\vec{\mathbf{r}},\omega)\,\frac{d\vec{\mathbf{r}}}{ds} \right] = \nabla n(\vec{\mathbf{r}},\omega) \qquad \frac{d^2 x}{dz^2} = \frac{1}{n(x,\omega)}\,\frac{d}{dx}\,n(x,\omega) \qquad \text{[II-11a]}$$

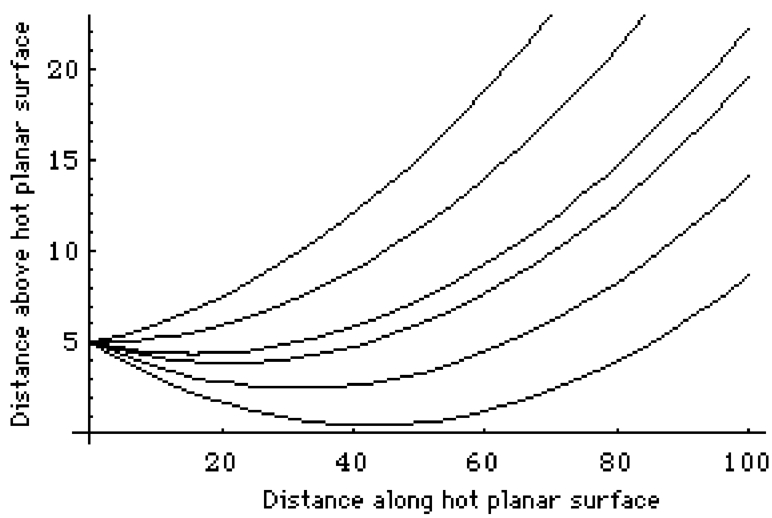
or from Equation [II-10]

$$\frac{d^2 x}{dz^2} \approx \kappa \,. \qquad \text{[II-11b]}$$

Thus, the ray trajectory is given by

$$\vec{\mathbf{r}}(z) = \left[\frac{\kappa}{2}\, z^2 + \tan\theta_0\, z + x_0 \right] \hat{\mathbf{x}} + z\, \hat{\mathbf{z}} \qquad \text{[II-12]}$$

Ray trajectories diverted by a hot surface



SECOND APPLICATION OF THE EIKONAL EQUATION: THE "ABCD" RAY MATRICES - A SYSTEMS APPROACH TO OPTICS

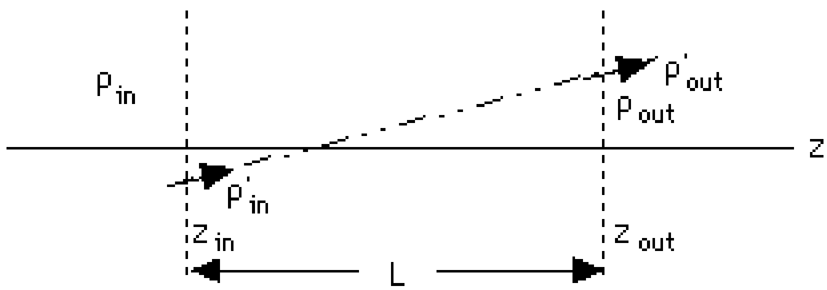
1. Uniform dielectric medium -- *i.e.* $n(\vec{\mathbf{r}},\omega)$ is a constant so that $\frac{d}{ds}\left[\frac{d\vec{\mathbf{r}}}{ds}\right] = 0$.

   Thus, the ray **must be a straight line** which may be written $\vec{\mathbf{r}} = s\,\vec{\mathbf{a}} + \vec{\mathbf{b}}$.

   In the two-dimensional paraxial approximation, we assume that $s \approx z$ and write

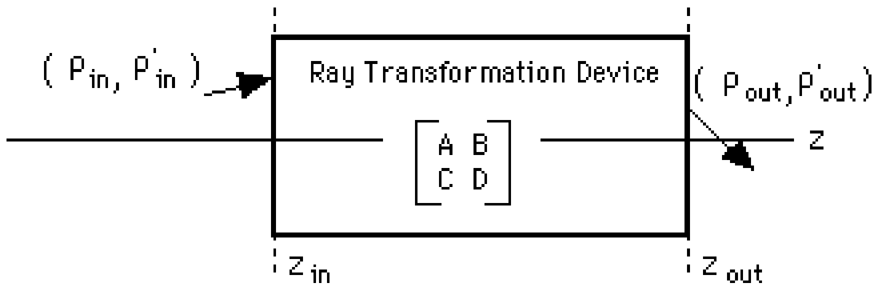
$$\rho_{out} = \rho_{in} + L\left.\frac{d\rho}{dz}\right|_{in} \quad \text{and} \quad \left.\frac{d\rho}{dz}\right|_{out} = \left.\frac{d\rho}{dz}\right|_{in} \qquad [\text{ II-13 }]$$

   where $\vec{\rho} = \rho\,\hat{\rho} = x\,\hat{\mathbf{x}} + y\,\hat{\mathbf{y}}$.



We may write results of this sort in the form of the famous and highly useful **ray transform** or **ABCD matrix** -- *viz.*

$$\begin{bmatrix}\rho_{out}\\ \rho'_{out}\end{bmatrix} = \begin{bmatrix}A & B\\ C & D\end{bmatrix}\begin{bmatrix}\rho_{in}\\ \rho'_{in}\end{bmatrix} \qquad [\text{ II-14 }]$$



In the case of a uniform dielectric

$$\begin{bmatrix} \rho_{\text{out}} \\ \rho'_{\text{out}} \end{bmatrix} = \begin{bmatrix} 1 & L \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \rho_{\text{in}} \\ \rho'_{\text{in}} \end{bmatrix} \qquad \text{[II-15]}$$

so that A = 1, B = L, C = 0, and D = 1

2. A dielectric discontinuity: Starting with Equation [II-7] and noting, once again, that *curl grad* $\{ \qquad \} = \vec{\nabla} \times \vec{\nabla} \{ \qquad \} = 0$ we see that

$$\vec{\nabla} \times \vec{\nabla} S(\vec{\mathbf{r}},\omega) = \vec{\nabla} \times \left\{ n(\vec{\mathbf{r}},\omega)\, \hat{\mathbf{s}} \right\} = 0 \qquad \text{[II-16]}$$

which is identical to the *saltus* condition on the electric and magnetic fields at a dielectric interface! Hence $\left\{ n(\vec{\mathbf{r}},\omega)\, \hat{\mathbf{s}} \right\}_{\text{tangent}}$ is continuous across the dielectric boundary so that $n_1 \sin\theta_1 = n_1 \sin\theta_1$ -- *i.e.* Snell's law! This result in the paraxial approximation (*i.e.*, $\sin\theta \approx \tan\theta \approx \theta$) may be written in ray matrix form as

$$\begin{bmatrix} \rho_{\text{out}} \\ \rho'_{\text{out}} \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} \rho_{\text{in}} \\ \rho'_{\text{in}} \end{bmatrix} \approx \begin{bmatrix} 1 & 0 \\ 0 & n_{\text{in}}/n_{\text{out}} \end{bmatrix} \begin{bmatrix} \rho_{\text{in}} \\ \rho'_{\text{in}} \end{bmatrix} \qquad \text{[II-17]}$$

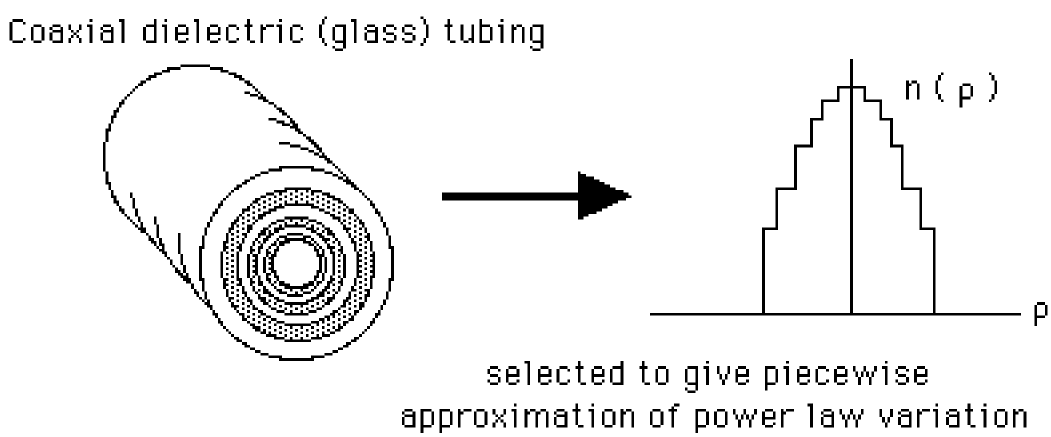
3. A "Thin" lenses: In passing we note that the ray matrix of a thin lens is given by or, perhaps more accurately, a thin lens is essentially defined by

$$\begin{bmatrix} \rho_{\text{out}} \\ \rho'_{\text{out}} \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} \rho_{\text{in}} \\ \rho'_{\text{in}} \end{bmatrix} \approx \begin{bmatrix} 1 & 0 \\ -f^{-1} & 1 \end{bmatrix} \begin{bmatrix} \rho_{\text{in}} \\ \rho'_{\text{in}} \end{bmatrix} \qquad \text{[II-18]}$$

4. Axially symmetric GRIN media: Consider the use of GRaded INdex technology to obtain an axially symmetric variation in the index of refraction of the form [9]

$$n(\rho) = n_{\mathrm{M}} \left[1 - \Delta \left(\frac{\rho}{a} \right)^{m} \right] \qquad \text{[II-19]}$$

9 A note on GRIN technology: In GRIN technology one builds up a glass rod with a specific radial index of refraction distribution by fusing a sequence of coaxially arranged glass tubes with appropriate index and diameter as illustrated in the following:



Within such a GRIN rod, we write $\vec{\mathbf{r}} = \rho\,\hat{\rho} + z\,\hat{\mathbf{z}}$ for the ray coordinates and $\vec{\nabla} n(\vec{\mathbf{r}}) = \hat{\rho}\,\frac{d}{d\rho}\,n(\rho)$ for the index variation. Using the eikonal equation -- *i.e.* Equation [II-9] -- in the paraxial approximation, we find

$$\vec{\nabla} n(\rho) = \frac{d}{ds}\left[n(\rho)\,\frac{d\vec{\mathbf{r}}}{ds}\right] \approx \frac{d}{dz}\left[n(\rho)\,\frac{d\vec{\mathbf{r}}}{dz}\right] = \frac{d}{dz}\left\{n(\rho)\left[\frac{d\rho}{dz}\,\hat{\rho} + \hat{\mathbf{z}}\right]\right\} \qquad \text{[II-20a]}$$

or

$$\frac{d}{d\rho}\,n(\rho) \approx \frac{d}{dz}\left[n(\rho)\,\frac{d\rho}{dz}\right] \qquad \text{[II-20b]}$$

Therefore

$$\frac{d^2}{dz^2}\,\rho \approx \frac{1}{n(\rho)}\,\frac{d}{d\rho}\,n(\rho) = \frac{d}{d\rho}\,\ln\left[n(\rho)\right] \qquad \text{[II-20c]}$$

or

$$\begin{aligned} \frac{d^2}{dz^2}\,\rho &\approx \frac{1}{n(\rho)}\,\frac{d}{d\rho}\left\{n_{\mathrm{M}}\left[1 - \Delta\left(\frac{\rho}{a}\right)^m\right]\right\} = \frac{1}{n(\rho)}\left[-n_{\mathrm{M}}\,\frac{m\Delta}{a}\left(\frac{\rho}{a}\right)^{m-1}\right] \\ &\approx -\frac{m\Delta}{a}\left(\frac{\rho}{a}\right)^{m-1} \end{aligned} \qquad \text{[II-20d]}$$

Doubtless, the simplest and most valuable instance is m = 2 -- *i.e.* what is usually called **parabolic** or **quadratic** material -- wherein

$$\frac{d^2}{dz^2}\,\rho \approx -\frac{2\Delta}{a}\left(\frac{\rho}{a}\right)^{2-1} = -\frac{2\Delta}{a^2}\,\rho = -\kappa^2\,\rho \qquad \text{[II-21]}$$

so that

$$\rho(z) = \rho_{\mathrm{in}}\,\cos(\kappa z) + \frac{\rho'_{\mathrm{in}}}{\kappa}\,\sin(\kappa z) \qquad \text{[II-22a]}$$

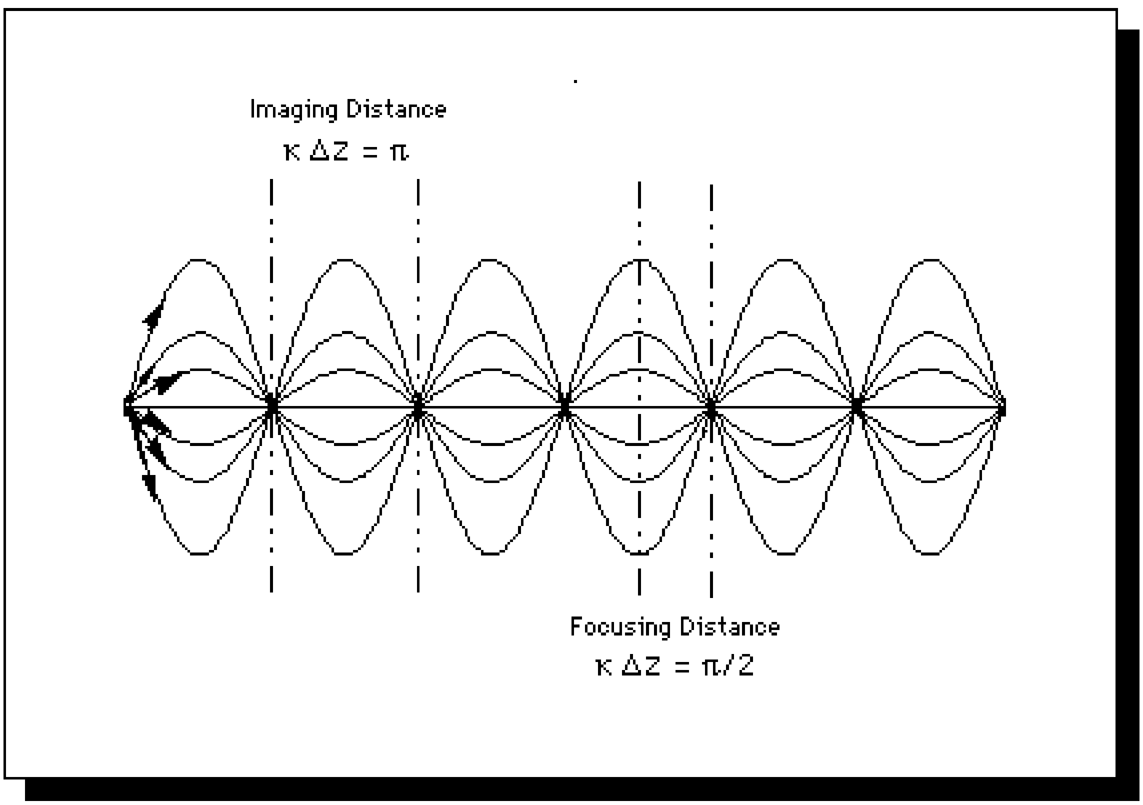
$$\rho'(z) = -\rho_{in}\,\kappa\,\sin(\kappa z) + \rho'_{in}\cos(\kappa z) \qquad \text{[II-22b]}$$

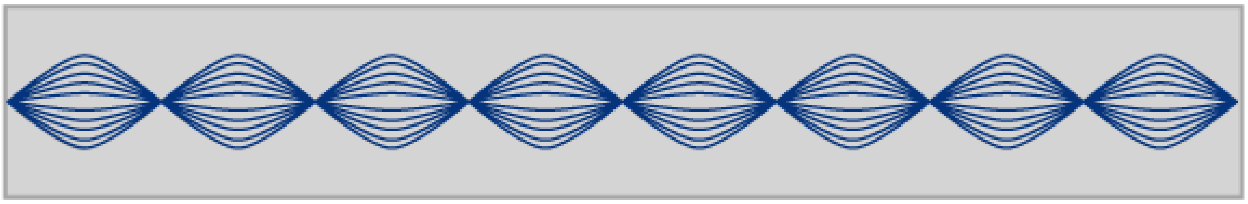
In terms of a ray transform matrix

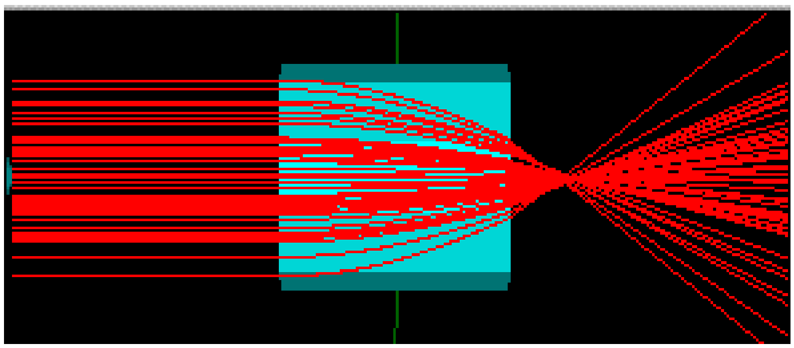
$$\begin{bmatrix} \rho_{out} \\ \rho'_{out} \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} \rho_{in} \\ \rho'_{in} \end{bmatrix} \approx \begin{bmatrix} \cos(\kappa z) & \kappa^{-1}\sin(\kappa z) \\ -\kappa\sin(\kappa z) & \cos(\kappa z) \end{bmatrix} \begin{bmatrix} \rho_{in} \\ \rho'_{in} \end{bmatrix} \qquad \text{[II-23]}$$

where $\kappa = \sqrt{2\Delta/a^2}$.

Ray trajectories confined in a GRIN rod.







ALTERNATIVE (HAMILTONIAN) DERIVATION OF EIKONAL EQUATION:

FERMAT'S PRINCIPLE [10]

Like most laws of physics, the equations of geometric optics can be derived from a **variation principle**. In this context the variation principle is called the Fermat principle which states that a ray always chooses a trajectory that minimizes[11] the optical path length -- *viz*.

$$\int_{P_1}^{P_2} n\left(x,y,z\right)\, ds = \text{minimum} \qquad [\text{ II-24 }]$$

---

10 See, for example, Dietrich Marcuse, *Light Transmission Optics*, Van Nostrand Reinhold (1972).

11 More precisely, the path must be a local *extremum* and in rare cases may, in fact, be a maximum. See R. Y. Luneberg, *Mathematical Theory of Optics*, University of California Press, Berkeley and Los Angeles (1964).

where the line element, ds, is measured along a ray and the two end-points P_1 and P_2 are fixed in space.[12] Analysis of the variation problem is facilitated by choosing the projected coordinate z as the new variable of integration. Accordingly,

$$ds = \sqrt{dx^2 + dy^2 + dz^2} = \sqrt{1 + x'^2 + y'^2}\; dz \quad , \qquad \text{[II-25]}$$

where $x' = \dfrac{dx}{dz}$ and $y' = \dfrac{dy}{dz}$, Fermat's variation principle is transformed into the more familiar **Lagrangian form** -- *viz.*

$$\int_{P_1}^{P_2} L(x, y, x', y')\, dz = \text{minimum} \qquad \text{[II-26]}$$

where

$$L(x, y, x', y') = n(x, y, z)\,\sqrt{1 + x'^2 + y'^2}\ . \qquad \text{[II-27]}$$

The minimization procedure is then well-known in the ***variational calculus*** and leads to the famous **Euler-Lagrangian equations** -- *i.e.*

$$\frac{d}{dz}\frac{\partial L}{\partial x'} - \frac{\partial L}{\partial x} = 0 \qquad \text{[II-28a]}$$

$$\frac{d}{dz}\frac{\partial L}{\partial y'} - \frac{\partial L}{\partial y} = 0 \qquad \text{[II-28b]}$$

When applied to the **Fermat Lagrangian**, as defined in Equation [II-27], these equations yield

---

12 From Equation [II-7] we see that

$$S(P_2) - S(P_1) = \int_{P_1}^{P_2} n\,(x,y,z)\; ds\ .$$

$$\frac{d}{dz}\frac{n x'}{\sqrt{1+x'^2+y'^2}} = \sqrt{1+x'^2+y'^2}\,\frac{\partial n}{\partial x} \qquad \text{[II-29a]}$$

$$\frac{d}{dz}\frac{n y'}{\sqrt{1+x'^2+y'^2}} = \sqrt{1+x'^2+y'^2}\,\frac{\partial n}{\partial y} \ . \qquad \text{[II-29b]}$$

Using Equation [II-25] we see that the **Euler-Lagrangian** equations may be expressed in the vector form as

$$\left\{ \frac{d}{ds}\left[n\frac{dx}{ds} \right] \frac{d}{ds}\left[n\frac{dy}{ds} \right] \right\} = \left\{ \frac{\partial n}{\partial x}, \frac{\partial n}{\partial y} \right\} \qquad \text{[II-30]}$$

which is precisely the content of Equation [II-9] -- QED.

HAMILTONIAN FORMULATION OF RAY OPTICS

The analogy between ray optics and particle mechanics is most striking when the equations of ray optics are expressed in Hamiltonian form.[13] To that end, we define the **generalized momentum** which is canonically conjugate to x and y by the vector equation

$$\left\{ p_x, p_y \right\} = \left\{ \frac{\partial L}{\partial x}, \frac{\partial L}{\partial y} \right\} . \qquad \text{[II-31]}$$

The Hamiltonian is then define in terms of the generalized momentum by the relation

$$H\left(x,y,p_x,p_y\right) \equiv p_x x' + p_y y' - L\left(x,y,x',y'\right) \ . \qquad \text{[II-32]}$$

With the assumed functional dependence of the Hamiltonian, we form the derivatives

13 The formal theory of optical systems was developed by Sir W. R. Hamilton in 1828-37.

$$\frac{\partial H}{\partial p_x} = x' + p_x \frac{\partial x'}{\partial p_x} + p_y \frac{\partial y'}{\partial p_x} - \frac{\partial L}{\partial x'}\frac{\partial x'}{\partial p_x} - \frac{\partial L}{\partial y'}\frac{\partial y'}{\partial p_x} \qquad \text{[II-33a]}$$

$$\frac{\partial H}{\partial p_y} = p_x \frac{\partial x'}{\partial p_y} + y' + p_y \frac{\partial y'}{\partial p_y} - \frac{\partial L}{\partial x'}\frac{\partial x'}{\partial p_y} - \frac{\partial L}{\partial y'}\frac{\partial y'}{\partial p_y} \; . \qquad \text{[II-33b]}$$

Given the definitional relationships embodied in Equation [II-31] we see that these expression reduce to one set of Hamilton's equation -- *viz*.

$$\left\{ \frac{dx}{dz}, \frac{dy}{dz} \right\} = \left\{ \frac{\partial H}{\partial p_x}, \frac{\partial H}{\partial p_y} \right\} . \qquad \text{[II-34]}$$

The other set of Hamilton's equation -- *viz*.

$$\left\{ \frac{d\,p_x}{dz}, \frac{d\,p_y}{dz} \right\} = \left\{ -\frac{\partial H}{\partial x}, -\frac{\partial H}{\partial y} \right\} . \qquad \text{[II-35]}$$

follow directly from the Euler-Lagrangian equations -- *i.e.* Equations [II-28a] and [II-28b] -- and the definitions embodied in Equation [II-31]. Using the Fermat Lagrangian we see that

$$\left\{ p_x, p_y \right\} = \left\{ \frac{\partial L}{\partial x}, \frac{\partial L}{\partial y} \right\} = \left\{ \frac{n\,x'}{\sqrt{1 + x'^2 + y'^2}}, \frac{n\,y'}{\sqrt{1 + x'^2 + y'^2}} \right\} \qquad \text{[II-36]}$$

and consequently that we may solve for $\left\{ x', y' \right\}$ in terms of $\left\{ p_x, p_y \right\}$ as

$$\left\{ x', y' \right\} = \left\{ \frac{p_x}{\sqrt{n^2 - p_x^2 - p_y^2}}, \frac{p_y}{\sqrt{n^2 - p_x^2 - p_y^2}} \right\} \qquad \text{[II-37]}$$

Substituting into Equation [II-32], we find an expression for the **Fermat** or **ray optics Hamiltonian** -- *viz.*

$$H = -\sqrt{n^2 - p_x^2 - p_y^2}\ . \qquad [\text{ II-38 }]$$

which resembles the mechanical Hamilton of a relativistic particle -- *i.e.*,

$$c\sqrt{m_0^2\, c^2 + p_x^2 + p_y^2 + p_z^2}\ .$$

But the analogy is even stronger in the paraxial approximation where the Hamiltonian is approximated by an expression which is identical in form with the Hamiltonian of a non-relativistic particle -- *viz.*

$$H = -n\sqrt{1 - \frac{p_x^2 + p_y^2}{n^2}} \approx \frac{p_x^2 + p_y^2}{2\langle n\rangle} - n \qquad [\text{ II-39 }]$$

when p_x and $p_y << \langle n\rangle$. [14]

[14] Applying the quantization rules of quantum mechanics to these Hamiltonians, we can go full circle and recover wave optics from ray optics. Equation [II-38] leads directly to the equivalent of the relativistic Klein-Gordon equation while the equivalent of the nonrelativistic Schrödinger equation follows directly from Equation [II-39].